An Invitation to Image Analysis

Kevin Vixie, Marian Anghel, and Brendt Wohlberg

Los Alamos National Laboratory

Abstract

The upcoming workshop *Image Analysis and Understanding Data from Scientific Experiments* at Los Alamos National Laboratory has two primary goals: to familiarize leading members of the image analysis community with problems of interest at LANL, and to introduce LANL scientists, many of whom have little expertise in image analysis, to some of the important tools and cutting edge developments in the analysis of image data.

These notes are intended to give examples of experimental data of interest, introduce some of the current techniques, and present our views and perspectives. We include three appendices covering a little more detail on the mathematical methods of image analysis, an annotated bibliography with links, and a discussion of a viewpoint which we believe is critical for the best results in the analysis of scientific data.

Contents

1 Introduction

The upcoming workshop *Image Analysis and Understanding Data from Scientific Experiments* at Los Alamos National Laboratory has two primary goals: to familiarize leading members of the image analysis community with problems of interest at LANL, and to introduce LANL scientists, many of whom have little expertise in image analysis, to some of the important tools and cutting edge developments in the analysis of image data.

We have written these notes in order to facilitate and inspire staff members with an interest in or need for image analysis tools. We show a few examples of data and techniques, explain our high level views and perspectives, and provide references and links to pertinent documents and Internet sites.

While the workshop has as its goal the use of image analysis techniques for the understanding of scientific data, (and there is certainly current research focused on this goal) we believe that other, apparently unrelated, research areas in image analysis are also potentially useful in pursuit of this goal.

2 Data and Problems of Interest at LANL

Here we give an idea of the experimental data of interest by briefly presenting a few examples. It should be kept in mind that this is really just the tip of the iceberg.

Proton radiography of high explosive experiments The data is generated by an operator that can be well approximated (after some preprocessing of the data) by a linear operator with a very large null space (many different objects give identical measurements). There are also noise and systematic errors, and the final goal is the understanding, on the basis of the data measured, of some physical quantity which is far from directly observable in the measured radiographic images.

Code validation with image data from shocked gas curtains In this case one has videos of fluid (a gas) evolving due to a shock wave hitting a gas curtain. The gas is illuminated by a "sheet" of laser light. The data is of high enough resolution to permit detailed comparisons with code.

Validation of thin film fluid equations from interferometric images In the experiments, a very thin fluid layer is driven up an inclined plane by a thermal gradient. The rates are on the order of mm/hour, and the evolution of the thickness of the film and the location of edges are measured interferometrically. The goal is the validation of computational models for thin film dynamics.(See Andrea Bertozzi's link in Section B.2 below.)

Identification of objects in hyperspectral satellite imagery Satellite and airborne sensors of increasing sophistication are providing images of increasing resolution, and with larger numbers of spectral bands. The ability to detect and identify spectral signatures and identify shapes in conjunction with those signatures has created many interested customers in this kind of inference from images. Examples include detection of plumes, identification of plume content, and detection of vehicle types and numbers from space.

Recognition of large dense objects or voids in the same Using background muon radiation, we may form tomographic radiographs by accumulating individual muon paths. These muon radiographs are then used to probe for abnormally density levels.

Recognition of faces in 2D images under varying poses Face recognition algorithms currently achieve greatly varying rates of recognition, depending, for the most part, on the the control of factors such as pose, lighting, and facial expression. The data has high enough dimensionality that some care with computational issues is required. While an enormous amount of work has been put into this area, we are a long way from any sort of definitive solution.

The brief, far from exhaustive list above begins to give an idea of the diversity of experiments, data, and questions leading to image analysis problems in which the goal *is not* the reconstruction of an image that is pleasing to the eye.

3 Examples of Current Techniques

Current techniques on the cutting edge include:

- Variational methods and PDE methods: total variation, active contours, level set methods, anisotropic diffusion, etc. Very briefly (see Section A.1) these methods allow the use of powerful insights and computational techniques well established for other purposes, which when brought to bear upon the image analysis tasks are found quite powerful.

- Markov random fields (Gibbs fields) and their extensions/generalizations: FRAME, Hierarchical MRF's, multiscale MRF's. These methods allow principled stochastic modeling of images whereby one often ends up with a maximum a posteriori solution.
- Multiresolution/wavelet methods and the use of custom basis methods which implicitly model noise, deal with nonlocality, and exploit efficient representation methods.

Recommended for the breadth of topics breezed through are the slides from a talk by Jackie Shen:

http://www.math.umn.edu/~jhshen/Venus/MichTour.pdf

We give more detail on the above techniques in Section A below. More thorough references and links are given there.

We close this section with examples of a few of the techniques of interest through the use of links to sites with pictures, movies, demos, etc.

Inpainting A Science article and some nice demos:

http://www.sciencenews.org/20020511/bob10.asp
http://mountains.ece.umn.edu/~guille/inpainting.htm
http://www.ece.umn.edu/users/marcelo/restoration.html

Total variation regularization Denoising results can be quite impressive:

http://www.math.ucla.edu/~imagers/htmls/rec.html

Level set methods The method introduced by Osher and Sethian with great success:

http://math.berkeley.edu/~sethian/level_set.html
http://www.math.ucla.edu/~imagers/htmls/lev.html

Active contours Originally introduced by Kass-Witkin-Terzopoulos:

http://www-sop.inria.fr/robotvis/const/ra/00/SuiviLS/
http://www.markschulze.net/snakes/

Dimension reduction PCA, POD, EOF, KH, and nonlinear generalizations, here we show that even PCA introduced early in the last century, is not dead:

http://www-white.media.mit.edu/vismod/demos/facerec/basic.html

Wavelet Methods Wavelet methods having to do with images are extraordinarily numerous. Here are a couple of nice examples and an irresistible demo on 1-D wavelets.

http://www-dsp.rice.edu/software/whmt.shtml http://www-dsp.rice.edu/software/ward.shtml http://cm.bell-labs.com/cm/ms/who/wim/cascade/math.html

Stochastic Modeling Most of these examples deal with Markov random fields:

http://www.utia.cas.cz/user_data/haindl/virtuous/dema/demtexsyn.html
http://www.ercim.org/publication/Ercim_News/enw44/haindl2.html
http://www.utia.cas.cz/user_data/haindl/virtuous/dema/dema.html
http://www.iro.umontreal.ca/~mignotte/postdoc.html
http://www.wi.leidenuniv.nl/home/lim/face.detection.html
http://rfv.insa-lyon.fr/~wolf/

Other In color images, correlations can be exploited:

http://www-sop.inria.fr/robotvis/personnel/David.Tschumperle/demos/color_restoration/

4 The Importance of Context

In the following section we argue that the complete context of an image is important when thinking about images and the information they represent. In particular, we argue that the whole flow of information from object to the perception/measurement, through the cognition/inference should be kept in mind when attempting to understand images.

What is an image? Here are typical definitions which, when taken as the domain of discussion for image analysis, are unnecessarily limiting in focus and scope:

Function: An image is a function $z = i(x, y)$ from the unit square, $[0, 1] \times [0, 1]$ to the unit interval $[0, 1]$.

Measurement: An image is a vector of length nm (or an $n \times m$ matrix) each entry of the vector being an 8, 12 or 16 bit binary number, etc.

Perception: An image is what I see with my eyes (this is particularly difficult to work with).

Images are important or interesting because they are measurements of something – typically 3-dimensional, and often dynamic – and so one should be thinking of the object, the measurement, and the interpretation instead of merely, for example, a vector of gray levels.

Thinking about measurements and how they arise leads us to a sequence of components: a possibly dynamic object that emits and/or reflects light, a medium that transmits and transforms that light, a recording system which reduces something infinite dimensional to a finite dimensional measurement, the presence of "noise" (making the whole process much more interesting and challenging) and finally, the perceiver/interpreter of the measurement, typically with a narrowly focused interest in the object which was measured in the image or sequence of images.

Keeping this sequence of components and their relationships in mind permits the use of the simpler definitions (e.g. an image is a vector) without the loss of depth risked by the use of the simplifications.

We do this now and define an image to be an finite or infinite dimensional vector formed through some (typically optical) measurement system, *but with the more extensive definition or context in mind, we consider the analysis of images to be divided somewhat naturally into three parts*:

Part 1 Image Formation (Object $\Rightarrow$ Image)

Part 2 Traditional Image Analysis (Image $\Rightarrow$ Image)

Part 3 Image Understanding (Image $\Rightarrow$ Semantic/Scientific Information)

One might argue that much of what falls under (2) above is actually very low level image understanding and that (2) and (3) are different only in degree, not quality. In our opinion, that this is not very useful. In fact, though (2) and (3) are very tightly coupled, it is still the case that segmentation or texture analysis is quite different than recognition of faces or the inference of elastic parameters from measured experimental images. But it is certainly the case that both (1) and (2) are closely involved in the pursuit of (3).

Appendices

A Image Analysis Tools: a Few More Details

In image analysis, there are three main approaches.

- The task may be formulated as a PDE or variational problem, the solution of which is the corresponding restored, segmented, denoised, etc. image. One can combine these techniques with multiresolution or hierarchical approaches to get an end result that moves closer to the image understanding end of things.
- The use of various multiresolution wavelet or wavelet-like representations to accomplish denoising, efficient modeling of images and their statistics (enabling e.g. more efficient object recognition).
- Stochastic methods handle noise naturally, yield estimates with explicit uncertainty quantification, and permit the flexible incorporation of prior information into image models. In this unbelievably brief glance, we will focus on a particularly successful method, that of Markov random field modeling.

Each of these approaches has something to teach us about images and the modeling of images. It seems likely that we will see more and more ties between these different approaches as each, in isolation, meets its limitations.

A.1 PDE and variational methods

The use of PDE and variational methods in image analysis has grown from a number of contributions in the 1980s and early 1990s (see for example Mumford-Shah [15, 16], Perona-Malik [19], Osher-Sethian [17], Rudin-Osher-Fatemi [20], Kass-Witkin-Terzopoulos [11]). Since then work in this area has exploded as experts in PDEs and variational methods have discovered image analysis and its rich source of inspiration and applications.

At the most basic level, the methods are based on one of two equation types:

$$u_t = F(u) \tag{1}$$

$$\hat{u} = \arg\min_u G(u) \tag{2}$$

where Equation 1 is an evolution equation in which the image u is evolving in time according to the PDE determined by the operator F, and Equation 2 tells us that the desired image u is in fact an optimal solution of the functional G. In the case of a time evolving PDE for an image u, we expect that the asymptotic solution is the "desired" restored, denoised, etc. image. A method of solution for the variational equation (Equation 2) is the method of time marching or gradient descent whereby one simply sets the negative functional or Frechet derivative of $G(u)$, $-DF(u)$, equal to u_t and letting this PDE evolve to the minimal point of F:

$$u_t = -DF(u). \tag{3}$$

(There are details: this method of solution is implicitly using the Riesz representation theorem to place u_t and $DF(u)$ in the same space.)

Before going to a specific list of methods we give a couple of links to papers by Chan, Shen, and Vese as well as a paper by Osher which goes into depth on important parts of the PDE/variational method spectrum:

- *http://www.math.ucla.edu/~imagers/htmls/internalreport/notice.ps*
- *ftp://ftp.math.ucla.edu/pub/camreport/cam02-43.ps.gz*

More specifically now, a number of methods whose various modifications and generalizations covers a great deal of ground in the PDE/variational world, are mentioned.

Inpainting In this image analysis enterprise one is attempting to restore images with missing pieces, scratches, etc. This is done in various ways by using the boundary information together with some regularization. One of the more interesting ways is the method which mimics fluid equations to "flow" the boundary data into the void. See the *http://www.ece.umn.edu/users/marcelo/restoration.html* for a nice demonstration of inpainting. The webpage at UCLA (*http://www.math.ucla.edu/~imagers/htmls/inp.html*) is also worth looking at.

Level set methods In this method curves in 2-D or surfaces in 3-D are thought of as level sets of functions $f : \mathbb{R}^2 \to \mathbb{R}$ or $f : \mathbb{R}^3 \to \mathbb{R}$ respectively. This has many advantages including: the PDEs governing curve or surface evolution are well behaved and changing topology of a curve or surface under evolution presents no special problems. The method, introduced by Osher and Sethian in 1988 [17] has been enormously successful: a Google search for level set method(s) yields a combined total of over 5700 hits!

Total variation minimization In total variation minimization [20] the image is denoised by minimizing the total variation $\int |\nabla f| \, dx \, dy$ under the constraint that the resultant restored image $u(x, y)$ is not too far (distance used to define too far usually the L_2 norm) from the measured, noisy image. In some cases the results are quite remarkable.

Mumford-Shah The Mumford-Shah functional [15, 16] minimizes the distance to data plus the amount of non-smoothness (not including the discontinuities) plus the length of the curves of discontinuities. The underlying theory is very beautiful as well as quite involved (see Ambrosio, Fusco, and Pallara).

Snakes, active contours, and geodesic flows Introduced in 1987, the Kass-Witkin--Terzopoulos model [11], in essence looks for a curve with minimal length, minimal curvature and which lies on an edge (image gray level discontinuity). This technique is used, for example, to find object edges in images.

The heat equation and more complicated cousins Denoising can be achieved via smoothing induced by an operator which uses the noisy image as initial data and simply "runs" the heat equation for some time. The loss of edge resolution is usually unacceptable and so one is lead to an approach introduced by Perona and Malik [19] in which the diffusion is anisotropic. In particular, the edges are much more nearly preserved. This method has a close relation to robust statistical estimation in which (very roughly) one uses only similar neighboring pixel values to compute the smoothing average.

A.2 Wavelets and other representations

Wavelet decompositions have, within the previous decade, become part of the standard mathematical toolbox for signal and image processing. The most significant success has been achieved in the field of image compression, where discrete wavelet transform based techniques such as EZW, SPIHT, and the recent JPEG2000 [25] standard offer considerable advantages over methods, such as the earlier JPEG standard, based on discrete Fourier and related transforms. Significant success has also been achieved in image denoising, where wavelet based techniques offer some of the most successful approaches. The range of applications is far more diverse, of course, including image watermarking, content-based retrieval, various pattern recognition tasks, stereo matching, and image fusion, to name a few.

In addition to wavelets, mention should be made of various adaptive transform strategies (some of which operate on overcomplete wavelet packet dictionaries) such as matching pursuit and basis pursuit, and the recent “beamlet”,“ridgelet”, and “curvelet” [23] decompositions developed by Donoho and collaborators.

A.3 Stochastic methods

The use of Markov random fields (MRFs) to model images was introduced in 1984 by Geman and Geman [7]. It is probably fair to say that MRF modeling and related techniques and extensions are the dominant influence in the stochastic modeling of images. This dominance is indicated by the more than 2400 citations of Geman and Geman.

Very roughly, a Markov Random Field is a spatial field of *sites* i (think of pixels) which are assigned *labels* $l(i)$ (think of gray levels) according to a probability distribution having the property that for any site i,

$$P(l(i) = \alpha \mid \text{all other sites}) = P(l(i) = \alpha | \text{neighboring sites of } i). \tag{4}$$

One can introduce prior knowledge through the specification of a neighborhood of influence for each site and the conditional distributions ($P(l(i) = \alpha|\text{neighboring sites of } i)$). The theorem of Hammersley and Clifford [9] (see also Besag [3]) showing that one may indeed specify the consistent conditional distributions via Gibbs potentials makes the specification of a MRF practical. (Another result (Griffeath [8]) permits the calculation of canonical potentials, which is of interest due to the non-uniqueness of potential functions – different potentials give rise to the same probability distribution.) One can build hierarchical and multiresolution MRF models, which permits, for example, the coupling of non-local effects into the neighborhoods.

An interesting extension and improvement to the MRF model is the FRAME model proposed by Zhu, Wu and Mumford in 1997 [36, 34, 35]. In this model, they use a sequence of filters to permit neighborhood influences which are more general than the usual ones, and in detail permit the constraints to be formulated in terms of matching statistics of training data but otherwise permitting maximum uncertainty in the probability model. More explicitly, one maximizes entropy under the constraint that the expected histograms of the outputs of a sequence of filters *matches* the corresponding average histograms from an ensemble of training data, one average histogram for each filter.

The key idea to the MRF model and its extensions is that one can build a consistent model for an image with embedded constraints in such a way that either Maximum Likelihood or, after selection of an appropriate prior, maximum a posteriori, can be applied to recover various estimates, restorations, and the like given the appropriate MRF model.

This leads to some challenging optimization problems which, after formulation of the correct model (i.e. one that correctly and consistent formalizes the constraints), are the main task to be pursued in MRF modeling of images.

One can find much more about MRF modeling and its extensions by looking at Kindermann and Snell [12] and Stan Z. Li [13].

B Annotated Bibliography and Links

The following annotated bibliography and list of links to groups doing exciting research is intended to provide some assistance in the significant effort required to familiarize oneself with a new research field.

B.1 Annotated bibliography

For an introduction to PDE and variational approaches to image analysis (and to a lesser extent image understanding) see the books by Osher and Fedkiw [18], Sapiro [21], Sethian [22], and Aubert and Kornprobst [2]. There is also the nice new book by Curt Vogel [26] which includes significant material relevant to the image formation problem. Stan Z. Li's book [13] on Markov Random Field modeling of images is the best monograph on this approach (which is probably the most established of the stochastic approaches to image analysis). As a companion to Li's book, we would highly recommend Kindermann and Snell [12]. More traditional image processing can be found, for example, in Jain's classic book [10]. For Wavelets, the best introductory references include those by Mallat [14] and Strang and Nguyen [24].

As a companion to the above books we would recommend Craig Evan's beautiful book on PDEs [5], Ziedler's books on applied and nonlinear functional analysis [32, 33, 29, 27, 31, 28, 30], Evans and Gariepy [6], and Ekeland and Temam [4].

Last, but not least, there is the wonderful book of Ambrosio, Fusco, and Pallara [1] which will be more than most need or desire, but is very highly recommended for those with a (some might say perverse) fascination with geometric measure theory.

B.2 Research groups and links

There are a multitude of investigators in image formation, image analysis, and image understanding. A list, by no means exhaustive, of significant investigators and research groups is given below.

UCLA The imaging group lead by Chan and Osher at UCLA has been very productive, especially in the area of PDE and variational methods for images. They also maintain a very nice library of reports which can be accessed at the following link:

http://www.math.ucla.edu/~imagers/

The entire collection of UCLA CAM reports includes work that is not image related, but is kept up to date better than the image groups list can be found at the following link:

http://www.math.ucla.edu/applied/cam/index.html

We also include a link to the UCLA vision lab run by Stefano Soatto:

http://www.vision.cs.ucla.edu/

Minnesota There is a growing effort at University of Minnesota spearheaded by Sapiro, Tannenbaum, Olver, and Shen, each of whose web pages follow:

http://www.ece.umn.edu/users/guille/
http://www.ece.umn.edu/users/tannenba/
http://www.math.umn.edu/~olver/
http://www.math.umn.edu/~jhshen/

Brown There are a number of notables in image analysis and pattern recognition at brown university including Mumford, Grenander, and one of the Geman brothers (Stuart). The list can be accessed through the pattern theory groups web page:

http://www.dam.brown.edu/ptg/

The other Geman (Donald) can also be accessed through the Brown site, but is at U Mass, Amherst and Johns Hopkins.

Caltech Perona's group at Caltech can be accessed through the following link:

http://www.vision.caltech.edu/.index.html

Carnegie Mellon Two webpages of interest at Carnegie of interest are the CIL (Calibrated Imaging Lab) webpage and the Computer Vision webpage:

http://www-2.cs.cmu.edu/~cil/
http://www-2.cs.cmu.edu/~cil/vision.html

Duke Andrea Bertozzi's group does research in image analysis as well as other areas such as microfluidics and swarming. Andrea and her collaborators are the source for the microfluidic data mentioned above. Included below is a link to her webpage and to an article in Science News Online concerning inpainting (also featured are the investigators Jackie Shen, Tony Chan, Guillermo Sapiro, and Marcelo Bertalmio):

http://www.math.duke.edu/~bertozzi/
http://www.sciencenews.org/20020511/bob10.asp

Yale Steven Zucker, who recently has recently become chairman of the Electrical Engineering Department at Yale, has a group working on early vision, grouping, and generic shape analysis. Below are his link (see his link "research") and a link to the Center for Computational Vision and Control:

http://www.cs.yale.edu/~vision/zucker/steve.html
http://cvc.yale.edu/frames.html

Courant Institute Eero Simoncelli works on various aspects of images including the statistical properties of natural images, visual motion and representation. His webpage has more details:

http://www.cns.nyu.edu/~eero/

SIAM Imaging Science Activity Group SIAM recently started another of the SIAM Activity Groups, this one in Imaging Science. The group organized a very successful conference in Boston in last spring (March 2002). Their links are as follows:

http://www.siam.org/siags/siagis.htm
http://www.math.ufl.edu/~bam/siagis/

There are many more See the the other links on the UCLA Imager's webpage as well as the growing list of links on the SIAM Imaging Science webpage.

C The Science Picture

Last, but certainly not least, we include a picture of the scientific enterprise that we believe is often missing from the perceptual framework of many investigators. We believe that this often leads to less than optimal results.

C.1 Representation, inversion, dynamics and metrics

Science attempts to provide a clearer picture of the world through a combination of modeling and experimentation. The goal of this enterprise is to understand, in as simple a manner as possible, the phenomena that are observable. This goal takes the form of the ability to control (think of jet aircraft, missile systems, and cures for diseases), to predict (think of the stock market and the weather), and to simplify to enable the understanding of yet more complex phenomena (think of the revolution brought about by Newton's work and the quantum mechanical work in the 20th century).

Science begins with observations, and through these observations one is inspired to take measurements of a phenomena which is very often modeled as a point in some finite or infinite dimensional state space undergoing temporal evolution governed by some set of principles/laws/equations. Recovering the state and/or the laws of evolution from a parameterized class, rejecting some hypothesis for the appropriate state space or governing equation, or inferring some unmeasured quantity from the measured ones captures a large number of the uses of scientific experiments and the data they represent.

To be able to arrive in the end at a defensible inference about the model/reality complex, one needs to have a consistent method for moving from data back to the model. This involves the following pieces which are often not given sufficient consideration.

Representation The object (state) and measurement spaces can be represented in many ways, all of which look equivalent until something practical, like computation or observation, is attempted. Practical efforts require truncations in expansions and the implied choice among possible finite dimensional approximations. Sometimes we are lucky enough to have a *natural* basis, distinguished by highly efficient representation (truncates well) , but usually we at least begin with something far from optimal. In addition to efficiency issues, the choice of representation can be used to implicitly ignore noise and invariances which one wants to factor out. Unfortunately, choice of representation is often left to whatever occurs first to the investigator, whatever was done before, or is easiest.

Inverse operator The analysis of scientific data (including image data) almost always involves inverse problems; the following therefore represent critical components in the generation of inferences from data:

- the inverse problem,
- its sensitivities,
- its null space or level sets (if they are nontrivial),
- the improvements possible through reparameterization, and
- the possibility for principled regularization.

Dynamics It is almost always the case that data is not time invariant, that is, there is variation in the temporal direction. In this case, we may use prior knowledge or hypotheses about the dynamics or class of dynamics to link data taken at different times together, effectively increasing the inferential power of the data. Parameter estimation such as Kalman filtering is just one example of this.

Metrics (and uncertainties) Finally, the comparison that generates an inference is often in the form of a metric which, far too often, is simply the "viewgraph norm" rather than some principled metric built through an understanding of the inverse, representation, and dynamics problems.

This brings us to the question of uncertainties.

C.2 Quantitative results and uncertainties

Understanding the representation, inverse, and dynamic problems helps us generate a metric which can be used to generate an inference. In some cases one can simply use the model to propagate the uncertainties forwards and/or backwards, possibly adding other uncertainties at the appropriate point in the flow of information. In other cases, one needs to rethink the equations of motion, perhaps using stochastic laws of evolution. The basic framework, though, remains the same — all the same issues must be dealt with. In both cases the final result might be a probability distribution over the state space, or an a posteriori distribution over the parameters of some model-class of dynamics given the data.

In the completely stochastic framework with stochastic states, dynamics, and measurements, the computations can be much more involved. Instead of moving points around some space X (often infinite dimensional), we must propagate functions (or measures) defined on X, around the space of such functions or measures. (Even the simpler case of involving deterministic dynamics to propagate uncertainty around is really moving a function forward rather than simply a point, but this task can actually be fairly straightforward.)

In the end, one must not ignore the uncertainties. The way the uncertainties are dealt with in the production of an inference needs to be thought about very carefully. We *do not* believe that there is some all encompassing way to do this, but we certainly advocate that the process be clear about assumptions and transparent in design.

References

[1] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford University Press, 2000.

[2] Gilles Aubert and Pierre Kornprobst. *Mathematical Problems in Image Processing*. Springer-Verlag, 2002.

[3] Julian Besag. Spatial interactions and the statistical analysis of lattice systems (with discussions). *Journal of the Royal Statistical Society, series B*, 36:192–236, 1974.

[4] Ivar Ekeland and Roger Temam. *Convex Analysis and Variational Problems*, volume 28 of *Classics in Applied Mathematics*. SIAM, 1999.

[5] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathmatical Society, 1998.

[6] Lawrence C. Evans and Ronald F. Gariepy. *Measure Theory and Fine Properties of Functions*. Studies in Advanced Mathematics. CRC Press, 1992. ISBN 0-8493-7157-0.

[7] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intellligence*, 6(6):721–741, 1984.

[8] D. Griffeath. Introduction to random fields. In J. G. Kemeny, J. L. Snell, and A. W. Knapp, editors, *Denumerable Markov Chains*, chapter 12, pages 425–458. Springer-Verlag, 2 edition, 1976.

[9] J. M. Hammersley and P. Clifford. Markov fields on finite graphs and lattices. unpublished.

[10] Anil K. Jain. *Fundamentals of digital image processing*. Pentice-Hall, 1989.

[11] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4), 1987.

[12] Ross Kindermann and J. Laurie Snell. *Markov Random Fields and Their Applications*, volume 1 of *Contemporary Mathematics*. American Mathematical Society, 1980.

[13] Stan Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer-Verlag, 2001.

[14] Stephane G. Mallat. *A wavelet tour of signal processing*. Academic Press, 1998.

[15] David Mumford and Jayant Shah. Boundary detection by minimizing functionals. In *IEEE conference on computer vision and pattern recognition*. San Francisco, 1985.

[16] David Mumford and Jayant Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Communications of Pure and Applied Mathematics*, 17:577–685, 1989.

[17] Stanley Osher and James A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.

[18] Stanley J. Osher and Ronald P. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer-Verlag, 2002.

[19] Pietro Perona and Jitendra Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE transactions on pattern analysis and machine intelligence*, 12(7):629–639, July 1990.

[20] Leonid Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268, November 1992.

[21] Guillermo Sapiro. *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press, 2001.

[22] James A. Sethian. *Level Set Methods and Fast Marching Methods : Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge University Press, 1999.

[23] J.-L. Starck, D. Donoho, and E. J. Candes. Very high quality image restoration by combining wavelets and curvelets. In *Wavelets: Applications in Signal and Image Processing IX*, volume 4478 of *Proceedings of the SPIE*, pages 9–19, San Diego, CA, USA, August 2001.

[24] Gilbert Strang and Truong Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge, 1997.

[25] D. S. Taubman and M. W. Marcellin. *JPEG2000: Image compression fundamentals, standards and practice*. Kluwer Academic Publishers, 2002.

[26] Curtis R. Vogel. *Computational Methods for Inverse Problems*. Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics, 2002. http://www.math.montana.edu/~vogel/.

[27] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications IIA: Linear Monotone Operators*, volume 2A. Springer Verlag, 1985.

[28] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications III: Variational Methods and Optimization*, volume 3. Springer Verlag, 1985.

[29] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications: Fixed Point Theorems*, volume 1. Springer Verlag, 1986.

[30] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications IV: Applications to Mathematical Physics*, volume 4. Springer-Verlag, 1988.

[31] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications IIB: Nonlinear Monotone Operators*, volume 2B. Springer Verlag, 1990.

[32] Eberhard Zeidler. *Applied Functional Analysis 1: Applications to Mathematical Physics*, volume 108 of *Applied Mathematical Sciences*. Springer Verlag, 1995.

[33] Eberhard Zeidler. *Applied Functional Analysis 2: Main Principles and Their Applications*, volume 109 of *Applied Mathematical Sciences*. Springer Verlag, 1995.

[34] Song Chun Zhu and David Mumford. Prior learning and gibbs reaction-diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11):1236–1250, 1997.

[35] Song Chun Zhu, Ying Nian Wu, and David Mumford. Frame: Filters, random field and maximum entropy: — towards a unified theory for texture modeling. *International Journal of Computer Vision*, 27(2):1–20, 1998.

[36] Song Chun Zhu, Yingnian Wu, and David Mumford. Minimax entropy principle and its applications to texture modeling. *Neural Computation*, 9(8):1627–1660, November 1997.